DOCUMENT RESUME

ED 246 392 CS 007 681

AUTHOR Nagy, William E.; And Others
TITLE Learning Words from Context. Technical Report No. 319.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jul 84
CONTRACT 400-81-0030
NOTE 64p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Basic Vocabulary; *Context Clues; *Learning Processes; Learning Theories; Reading Comprehension; Reading Habits; Reading Improvement; *Reading Research; *Reading Skills; *Vocabulary Development

ABSTRACT

To test the hypothesis that a large proportion of school children's vocabulary growth occurs through incidental learning from written context, a study was conducted using 57 eighth grade students of average and above average reading ability. The subjects read a 1,000 word expository or narrative text. After reading, subjects completed two vocabulary assessment tasks on 15 target words from each passage, an individual interview, and a multiple choice test, both designed to test partial knowledge of word meanings. Results of within-subject, hierarchical regression analyses showed small but statistically reliable gains in word knowledge from context. Tentative extrapolations from the results and current estimates of the volume of children's reading resulted in the conclusion that incidental learning from context accounts for a substantial proportion of the vocabulary growth that occurs during the school years. (Author/CRH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED246392

CENTER FOR THE STUDY OF READING

Technical Report No. 319

LEARNING WORDS FROM CONTEXT

William E. Nagy
Patricia A. Herman
Richard C. Anderson
P. David Pearson

University of Illinois at Urbana-Champaign

July 1984

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman, Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

The research reported herein was supported in part by the National Institute of Education under Contract No. 400-81-0030. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the office should be inferred. The authors wish to thank Mr. Charland, Principal, and the students and staff of Forest Trails Junior High School, Park Forest, Illinois, for participating in the study.

CS007681

EDITORIAL BOARD

William Nagy
Editor

R. Logan Bickford
Harry Blanchard
Pat Chrosniak
Avon Crismore
David Dunning
Linda Fielding
Patricia Herman
Asghar Iran-Nejad
Steve Levine
Margie Leys
Paul Mayberry
Mary Moes
Carol Peterman
Michael Reddix
Theresa Rogers
Judith Scott
Ileana Seda-Santana
Fernando Senior
Marcy Stein
Janice Stewart
Behrooz Tavakoli
Paul Wilson

Abstract

School children appear to increase their vocabularies by thousands of words per year. Many have hypothesized that a large proportion of this growth occurs through incidental learning from written context. However, experimental research has until now failed to provide unequivocal support for this hypothesis. The present study attempted to determine whether students do acquire measurable knowledge about unfamiliar words while reading natural text. Fifty-seven eighth grade students of average and above average reading ability read either an expository or a narrative text about 1000 words in length. After reading, subjects completed two vocabulary assessment tasks on 15 target words from each passage (thus serving as controls for the passage not read), an individual interview and a multiple-choice test, both designed to tap partial knowledge of word meanings. Results of within-subject, hierarchical regression analyses showed small but statistically reliable gains in word knowledge from context. Tentative extrapolations from the results and current estimates of the volume of children's reading lead us to believe that incidental learning from context accounts for a substantial proportion of the vocabulary growth that occurs during the school years.

Learning Words from Context

This paper represents one step in a program of research aimed at testing the following hypothesis: Incidental learning from context during free reading is the major mode of vocabulary acquisition during the school years, and the volume of experience with written language, interacting with reading comprehension ability, is the major determinant of vocabulary growth.

Incidental learning from context has traditionally been assumed to be one cause, if not the major cause, of vocabulary growth. Boettcher's (1980) dissertation quotes sources as far back as St. Augustine in support of this view. As stated somewhat more recently by Gray and Holmes (1938),

> [W]e know from experience that practically all pupils acquire many meanings from the context with little or no help from teachers (p. 28) . . . . Growth [in vocabulary] can be secured most effectively through wide silent reading with little or no guidance in the understanding or use of words. (p. 35)

On the other hand, strong experimental evidence for this position does not seem to be available. In a recent article, Jenkins, Stein, and Wysocki (in press) assert:

> We have been unable to locate any experiments conducted under relatively natural reading conditions which directly studied learning (as opposed to deriving) word meanings from

context. Such demonstrations are required to support the learning from context position, and to move it beyond its current status of a default argument.

The "default argument" for learning from context rests on the large and otherwise unexplained volume of vocabulary learning that goes on during a child's school years. Even by extremely conservative estimates, children learn upwards of 600 words per year during their school years. Some researchers (e.g., M. K. Smith, 1941; Templin, 1957) have reported children's vocabularies to increase by more than 5,000 words a year. Nagy and Anderson (1984) present evidence that the actual rate of vocabulary growth during school years is likely to be closer to these higher figures (see also Nagy & Herman, in preparation).

What is intriguing is that this massive vocabulary growth seems to occur without much help from teachers. Surveys of instruction (Durkin, 1979; Beck, McKeown, McCaslin, & Burkes, 1979; Jenkins & Dixon, 1983) show relatively little direct instruction in vocabulary taking place. How and where all this vocabulary learning occurs is still open to question. The only plausible explanation seems to be some type of incidental learning from context. However, the relative contribution of conversations with adults or peers, television, classroom discussion, school reading, or free reading is not known.

The puzzle is that previous research has failed to provide solid support for the hypothesis that learning from context is a

major source of vocabulary growth. Several studies have found learning from context to be ineffective when compared to other ways of acquiring new vocabulary. Other studies have reported successful learning from context; however, these studies have generally involved tasks which are inherently easier than learning from natural context during normal reading. Thus, they may have overestimated the efficacy of learning from context, and therefore do not provide a satisfactory basis for evaluating the role of incidental learning from context in children's vocabulary growth.

There are three major ways in which previous studies have been likely to overestimate learning from context. These are not necessarily flaws in the studies themselves, since the studies did not all have as their purpose evaluating the role of learning from context in overall vocabulary acquisition. However, in terms of the hypothesis we are considering, these constitute failures to achieve ecological validity.

First, some studies (e.g., Carroll & Drum, 1983; Sternberg & Powell, 1983) deal with subjects' ability to derive word meanings from context; that is, subjects are given explicit instructions to figure out the meaning of unfamiliar words with the text in front of them. Certainly the ability to do this is related to the ability to learn the meanings of new words from context. However, the percentage of word meanings that can be derived from context overestimates the percentage that would be learned during

normal reading. The chief reason is that in normal reading a person often skips over an unfamiliar word, rather than focusing more attention on it (Freebody & Anderson, 1983).

Second, many studies have investigated subjects' ability to learn meanings from unnaturally informative contexts. Some studies (e.g., Gipe, 1979) have used such rich contexts that they really measured subjects' ability to learn word meanings from definitions. Other studies, while avoiding this, have nevertheless used contexts much more informative than are found in most normal text (e.g., Jenkins et al., in press). Again, such studies overestimate the amount of learning from context that would occur in normal reading; many, probably most, contexts in normal text give little information about word meanings.

Third, as Jenkins and Dixon (1983) have pointed out, how easy it will be to learn a new word from context depends upon characteristics of the word and its associated concept. Most pertinent to the present discussion is the distinction they make between learning a new label for a familiar concept, and learning a new label for a new concept. Studies of learning from context frequently have focused only on the former task, either by using blanks or nonsense words to replace real, known words, or else by selecting difficult real words for which familiar synonyms exist (e.g., Rankin & Overholzer, 1969; Werner & Kaplan, 1952; cf. Boettcher, 1980, pp. 54-55). Learning a new label for a familiar concept, or figuring out which familiar concept fits into a slot

in text, will almost always be easier than learning both a new concept and a new label. Studies that look only at the easiest cases of learning from context give too optimistic a picture of the amount of learning from context that takes place in normal reading. Judging from examples of the words used, many studies of learning from context suffer from this limitation.

Previous studies of learning from context have generally had one or more of these weaknesses. To the extent that this is true, they overestimate learning from context in the normal reading situation; thus, whatever learning from context they do show does not constitute strong support for the hypothesis that learning from written context is a major factor in vocabulary growth.

Furthermore, several studies have shown learning of word meanings from written context to be a relatively ineffective process (e.g., Gibbons, 1940; Sachs, 1943), especially when compared with intensive direct instruction (Jenkins, Pany & Schreck, 1978; Johnson, Toms-Bronowski & Pittelman, in press; Margosein, Pascarella & Pflaum, 1982). This is true even for studies which might be expected to overestimate learning from context, because rich and informative contexts were used (e.g., Jenkins, Stein & Wysocki, in press). Even using extremely rich contexts, Gipe (1981) was unable to replicate the relative advantage of learning from context over alternative methods of vocabulary instruction which she had found in her earlier (1979)

research. The fact that even studies which might be expected to overestimate learning from context have found it to be relatively ineffective gives all the more grounds for questioning the importance of learning from written context.

Beck, McKeown, and McCaslin (1983) voice a general skepticism of learning from written context as the source of vocabulary growth:

> The point of our discussion has been that contexts occurring in text selections do not reliably assist readers in discovering the meaning of an unknown word. However, even the appearance of each target word in a strong, directive context is far from sufficient to develop full knowledge of word meaning . . . . The reliance of basal reading programs on story context and independent use of the glossary as the central methods of vocabulary development is at best appropriate for the most motivated and competent readers. Children most in need of vocabulary development, less-skilled readers who are unlikely to add to their vocabulary from outside sources, will receive little benefit from such indirect opportunities (pp. 180-181).

We cannot argue with the claim that for a given word the quickest way to impart thorough knowledge of its meaning is via direct instruction. We maintain, however, that the efficacy of learning from context must be evaluated, not in terms of short term competition with direct instruction, but in terms of the

volume of vocabulary growth that can be accounted for over an extended period of time. Previous research in learning from context has not provided a sufficient basis for this kind of evaluation. In the present study, we attempt to extrapolate from the short term results to calculate the proportion of total vocabulary growth that can be attributed to incidental learning of word meanings from written context.

The Incremental Nature of Word Learning

While there are studies which show that learning of word meanings from context can occur, the data seem to indicate that it is a rather ineffective process. Deighton (1959) lists some likely reasons for this: (a) Only some contexts, probably a small percentage, give much information about the meaning of a word, (b) at best, only one of the possibly many meanings of the word is supported by the context, and (c) the context will supply information about only some aspects of this one meaning of the word. Deighton concludes that vocabulary growth from context is a gradual matter.

Research in both vocabulary instruction and early vocabulary acquisition supports the idea that that learning individual word meanings is a gradual process. Boettcher (1980), Dale, O'Rourke, and Bamman (1971), and Eichholz and Barbe (1961) offer models of word learning which differ in details as to the number or nature of intermediate stages of knowledge, but all agree that word learning often proceeds by small increments.

A variety of psycholinguistic research shows that children initially have incomplete knowledge about the meanings of words (e.g., Clark, 1973; Gentner, 1975). While the exact interpretation of the data is not always clear (cf. Carey, 1982), it is apparent that children's first representation of the meaning of a word often overlaps only partially with that of an adult.

There is also evidence available to support the belief that substantial, if incomplete, knowledge about the meaning of a word can be gained through one or a small number of exposures. First, there is indirect evidence that children are learning words somehow at a remarkable rate. Statistical studies of word distribution (Carroll, Davies & Richman, 1971) show that the bulk of the words in the language are of low frequency; almost 70% of the words (types) that appear in printed school materials for grades three through nine occur once or less in a million words of text. If a child is learning the meanings of such words from context, it must be on the basis of very few encounters.

Second, there is direct evidence that children can and do gain substantial, if partial, knowledge of a word's meaning from a single encounter in context. In Carey's (1978) study, children were exposed to a new color word in fairly rich but natural contexts. Carey found that very few exposures to a new word were necessary for children to learn something about its meaning--in this case, at least that it was a color word. She concludes that

this first stage of lexical acquisition, "fast mapping," is a very efficient process, but that complete learning of a word's meaning is a gradual process, probably extending over years of time in which the word is encountered repeatedly.

We hypothesize, then, in agreement with Deighton (1959), that incidental learning from context proceeds in terms of small increments, so that any one encounter with a word in text will be likely to produce only a partial increase in knowledge of that word. On the other hand, we also hypothesize that learning from context is more effective than many have assumed. Although a single encounter with a word would seldom lead to a full knowledge of its meaning, we believe that substantial, if incomplete, knowledge about a word can be gained on the basis of even a single encounter. Therefore, if coupled with a sufficiently large volume of exposure to written language, incidental learning from context should be able to account for a substantial amount of vocabulary growth.

The failure of many studies to demonstrate appreciable learning from context, we would argue, lies in the insensitivity of the measures of word knowledge to small increments of learning. Often researchers have chosen words of very low frequency to insure that subjects have no prior knowledge of their meanings--but then test for learning from context in a way that requires full knowledge of the word's meaning for a correct answer. If learning from context normally proceeds in terms of

small increments, such an approach must substantially underestimate the amount of learning from context that goes on. In this study, on the other hand, we employed measures of word knowledge--both interviews and multiple choice questions--specifically designed to be sensitive to partial knowledge of word meanings. This was intended to enable us to detect the incidental learning of word meanings hypothesized to take place even in the not-especially-rich contexts found in natural text.

Method

Subjects

Seventy average and above average eighth grade students were identified by school personnel and by the Gates-MacGinitie reading test. The mean reading comprehension percentile was 71.5, s.d. = 17.4, range 28 to 99. The mean vocabulary percentile was 73.2, s.d. = 16.7, range 39 to 99. Out of the pool of 70 students, 63 took a checklist vocabulary test. Sixty students were present for the main study. Of these, complete data were available for only 57; two did not finish the multiple-choice test, and one was found not have taken the checklist test. Results are reported only for the 57 students for whom complete data are available.

Students were randomly assigned to read either a spy narrative or an exposition on river systems (see Materials), and to one of the versions of the vocabulary tasks. To assess the equivalence of the narrative and exposition groups, six

comparisons of pre-experimental knowledge and ability were made involving knowledge of target words from the narrative and expository passages, background knowledge relevant to each passage measured in terms of topic-related words not occurring in the passages, and standardized comprehension and vocabulary scores. No differences between the groups were found (all Fs < 1.0).

Materials

Texts. Two junior high level texts of different genre were chosen. One, "The Midnight Visitor" (Arthur, 1981) from the basal Beacons, was a mystery with about 1000 words. This narrative text was used verbatim. The other, taken from a chapter entitled "Water Systems" in Earth Science (Bishop, Sutherland & Lewis, 1981), was an expository text about 960 words long. One paragraph and a few sentences were deleted from it to insure that it would be a self-contained unit of approximately the same length as the other text, but no other changes were made. Although no systematic comparisons were made, both texts could be considered typical material for junior high students. Both texts were reproduced without illustrations.

Target vocabulary words. The fifteen most difficult words from each text were selected as target words. "Word" in this case includes both single words and two-word compounds such as suspended load and drainage basin. Two measures of difficulty were taken into account: Several raters with teaching experience

were asked to circle the most difficult words or phrases in the passages and the Standard Frequency Index from Carroll, Davies and Richman (1971) was considered. The final set of target words included those words identified as difficult by all raters, and those words identified as difficult by all but one rater that had the lowest frequencies. The target words were of low frequency with the exception of the two-word compounds (e.g., drainage basin), which have much lower frequencies as compounds than the frequency of either component, and frequent words which were used in the passage with less frequent meanings (e.g., bed = "riverbed" or divide = "a ridge or high ground separating drainage basins"). A list of the target words is given in Table 1.

Insert Table 1 about here.

The target words varied in several respects. Some were morphologically simple (twang, rill), others contained suffixes that might reveal something about their syntactic function (authentic, turbulent), and others were compounds whose parts might help in deducing their meanings (floodplain, suspended load). Some of these words constituted new labels for familiar concepts (e.g., espionage = "spying"), whereas others (e.g., drainage basin) presumably represented unfamiliar concepts.

The use of real words in natural texts increases ecological validity, but it makes it difficult to assure that subjects did

not already know the meanings. However, both the results of the checklist vocabulary test, administered several days before the main body of the study, and the performance of subjects on target words not in the passage read, served as statistical controls for the likelihood of a word having been known before the experiment. Also, the presence of some partially known words enabled us to investigate increase in knowledge of such words, an important aspect of vocabulary growth overlooked in previous studies.

Checklist vocabulary test. A checklist test was developed, using the guidelines suggested by Anderson and Freebody (1983), as a measure of the vocabulary knowledge of subjects prior to reading the experimental passages. In this test, a subject simply indicated whether or not the meaning of a word was known. Some of the items in the test were English-like nonwords; these provided the basis for a correction to adjust for guessing and response bias.

The checklist test was chosen because it gives the subject no information about the meanings of the words tested. It is also sensitive to partial word knowledge; subjects tend to mark a word as known if they have even a partial grasp of its meaning (Anderson & Freebody, 1983). A weakness of a checklist test is that it is not suitable for use as a pre- and post-test.

The checklist test used in this study consisted of 186 items, in the following categories:

(1) The target words from each of the two passages. (Compounds such as oxbow lake were divided into two words for the purpose of the checklist test, so there were a total of 19 target word items.)

(2) Fifteen background knowledge words for each passage; that is, fifteen words related to espionage (e.g., wiretap, surveillance) and fifteen related to river systems (e.g., aquifer, glacier), which did not occur in the passages.

(3) Thirty general vocabulary items, chosen to represent a range of difficulty.

(4) Thirty-two decoding distractors. These are items which would be marked as known only on the basis of a decoding error (e.g., weast, robbit).

(5) Thirty pseudo-derivatives. These are not existing words of English, but are constructed from existing English stems and affixes (e.g., successment, desertitude).

(6) Thirty nonwords. Items in this category (e.g., felinder, werpet) are also not existing words of English. Furthermore they do not belong to either of the two preceding categories. That is, they are not constructed from real English stems and suffixes, nor could they be mistaken for a real word if some plausible error were made in decoding. Only these nonwords were used in computing the correction factor for a subject. Four versions of the checklist test were constructed, each with a different ordering of items.

Story memory task. This task provided a delay between the reading of the passage and the interview about the meanings of the target words. While the task kept the subjects' attention on the passage read, it did not provide any additional information about the meanings of the target words.

Items in the task consisted of a word or two-word compound followed by the phrases "saw it in passage" and "have seen it elsewhere." Subjects were asked to put an X through either or both of these phrases if they applied. Four versions of this task were constructed, each with the items in a different order.

Multiple choice test. A multiple choice test for measuring degrees of knowledge of word meanings was developed. For each of the thirty target words, a concise definition was chosen to serve as the correct answer. For example, the short definition for divide was "a ridge or high ground separating areas belonging to two different river systems;" for envision it was "to imagine or picture something."

For each target word, test items were constructed at each of three levels of difficulty. An example of the three levels of difficulty for one of the target words is shown in Table 2.

Insert Table 2 about here.

Levels of difficulty were based on the similarity in meaning between the target word and the concepts represented by

the distractors. At the highest level of difficulty, distractors represented concepts similar to or closely associated with the meaning of the target word. At the lowest level of difficulty, distractors were chosen to be as dissimilar from the target word meaning as possible, even in terms of the implied part of speech. At the intermediate level of difficulty, distractors were chosen to be mostly in the same part of speech, but otherwise fairly diverse semantically.

As often as possible, at least one distractor was shared by adjacent levels of difficulty. For example, in the item in Table 2, the distractor "the illegal transportation of goods across a border" is used both at the lowest and intermediate levels of difficulty. This is to lessen the extent to which subjects could guess the correct answer simply by remembering which choices were common to all items for the same word.

Three types of distractors occurred in the items. First of all, the correct answers for target words were used as distractors for other items. At each level of difficulty, each target word's definition occurred at least once as a distractor in another item, but no more than three times. It was hoped that this repeated occurrence of the target word definitions would make it more difficult for subjects to pick up the association between the target word and its definition from the test alone.

In addition to the target word meanings, short definitions of other concepts in the experimental passages were used as

distractors. For example, one distractor was "material rolled along the bottom of a river channel by the current"--a definition of the concept bed load mentioned in the text. Especially at the highest level of difficulty, it was also necessary to use a third category of distractor, namely definitions of concepts closely related to or similar to the target word meaning which did not occur in the text.

Each multiple choice item contained the correct answer, four distractors, and a "don't know" option. Position of the correct answer was assigned in quasi-random fashion, with correct answers occurring with equal frequency in the first five positions, and in three different positions for any given target word. The "don't know" option was always in the last (sixth) position.

The multiple choice test was divided into three blocks, with each block containing one item for each target word. Level of difficulty and order were counterbalanced for the items. Each block was divided into two sub-blocks; target words were assigned to sub-blocks such that two items for the same target word never occurred in adjacent sub-blocks. Thus there were always at least 15 test items between any two appearances of the same target word. Order of items within the sub-blocks was randomized. Six versions of the test were constructed, with three different orders of the blocks and two different orders of sub-blocks within blocks.

Procedure

Three days before the main part of the study, the checklist vocabulary test was administered to the group of 70 eighth grade students. After a researcher had read the direction page aloud, students completed the test at their own pace.

The main part of study took place over a two day period during regular school hours. Although students knew they were in a university study, they did not know the purpose of the tasks. All work was monitored by one of the researchers to insure that students understood instructions for the tasks and worked individually.

After a group of 5-7 students arrived in the testing room, a researcher read a set of general introductory remarks. No mention of vocabulary or themes in the passages was made. Thus, care was taken to have the students read the text under as natural conditions as possible.

Following the introduction, the researcher passed out copies of the passages, alternating the two types of text between students. Answer booklets were distributed face down and the students in a session received different versions. Students were not allowed to open the booklets until directions were given.

Before reading the passage, a researcher read aloud the direction page preceding the text. Students were told they would have ten minutes to read their passage, could reread it as much

as they liked during that time, and would be asked questions about the passage without being able to see the text. Students who finished early and did not choose to reread or study the passage sat quietly. Students were not allowed to do other work or to talk. After ten minutes, all passages were collected.

Students then proceeded to the answer booklets. Directions for the story memory task were read to the students. Since no two students in the same session had the same version of answer booklet, the likelihood of successful copying was reduced considerably. Students were allowed to work at their own pace. Although finishing times varied, no student took more than 20 minutes to complete both the reading of the text and the story memory task.

Immediately after completing the story memory task, each student was assigned to one of several trained interviewers for individual interviewing on the meanings of the target words. Before the student's arrival, the interviewer had randomized the 30-card deck of target words by shuffling it. Then, with the student looking at the sample target words, the interviewer read the instructions detailing the task of defining target words. As students attempted to define the sample words, the interviewer used the same prompts as would be used later for the target words. Interviewers stressed the importance of sharing partial word knowledge, giving an example of such sharing with one of the difficult sample words.

When the student understood the task, the interview began. Holding up one of the 3 x 5 cards displaying a target word, the interviewer asked the student to say the word. Mispronunciations were not corrected. Next, the student was asked to tell what the word meant or to use it in a sentence. If a clear answer was given, the interviewer asked the next word. Interviewers had been trained beforehand on what the correct definitions were. If an unclear and/or incomplete definition was given, the interviewer used one of the following prompts depending on what the student had already said: (a) "That's part of the meaning. Can you make it more clear?" (b) "That's one meaning. Do you know another meaning for this word?" and (c) "Does this word remind you of anything?" (see Figure 1). Interviews lasted about 30 minutes.

Insert Figure 1 about here.

The last part of the procedure was the multiple choice test. Students worked through the test at their own pace, taking approximately 30-45 minutes to complete it. The researcher monitored each student to be sure the question numbers and answer sheet numbers matched.

Scoring

Interviews. Interviews were scored on a four-point scale by two raters who were blind as to which story a student had read.

To maintain consistency in scoring and to minimize any bias a rater could develop for a particular student's answers, raters scored all of the answers to one word before going on to the next word. Raters independently scored the interviews according to the following criteria: (a) zero points for an answer with no correct knowledge, (b) one point for an answer with minimal partial knowledge, that is a little more than nothing with at least some real, correct knowledge, (c) two points for an incomplete answer which displayed substantial correct knowledge, but was still missing some important component of meaning, and (d) three points for a totally correct answer. An example of scoring for the word disillusioned is given in Table 3.

Insert Table 3 about here.

Inter-rater reliability, measured in terms of how subjects ranked for number of words known at each level, was .72 for Level 1, .73 for Level 2, and .70 for Level 3. To maximize reliability, both raters scored all the interview data, and all disagreements were examined and resolved.

Results

The basic results of this study are presented in Table 4. It can be seen that at each level of difficulty, for both the interview and multiple choice test, a greater proportion of the

target words from a given passage were known by the subjects who had read that passage than by the subjects who had not.

Insert Table 4 about here.

Tables 5 and 6 summarize hierarchical multiple regression analyses that were performed following the logic of within-subjects analysis of variance. The comparisonwise alpha level was set at .01 to keep the experimentwise error rate within reasonable bounds. The dependent measure in both the interview and multiple choice analyses was whether or not a subject knew a given word at a given level, expressed as a percentage.

Insert Tables 5 and 6 about here.

In the interview analysis, the variables were entered in the following order: (a) Subject's Grand Mean, the subject's mean performance on all target words, entered first in the equation to remove variance associated with differences between subjects, (b) Prior Target Word Knowledge, the subject's reported prior knowledge of the specific target word, based on the pre-experimental checklist test, (c) Level, the level of the criterion for word knowledge (for example, if a subject's response in the interview was scored as reflecting level 2 word knowledge, the subject was counted as knowing the word for Level =

1 and Level = 2, but not for Level= 3), (d) Text Read, the text read by the subject, 1 for narrative and -1 for exposition, (e) Word Source, which identifies the text in which that particular target word occurred, (f) Learning from Context, (g) Reading Comprehension Ability, and (h) the Learning from Context by Reading Comprehension Ability interaction.

The analysis of the multiple choice data includes the same variables. Level, however, is defined slightly differently; in this case it simply represents the level of difficulty of a given multiple choice item. The multiple choice analysis also includes the variable Position, the position of the item in the multiple choice test.

Of primary concern is the variable Learning from Context. This variable actually is the interaction of Text Read and Word Source. It represents the degree to which subjects did better on words from the passage they read--that is, the extent to which they learned word meanings from context. In both analyses, Learning from Context was highly significant.

The interaction of learning from context with the standardized measure of reading comprehension was not quite significant in either the interview analysis ($.01 < \underline{p} < .05$) or the multiple choice analysis ($.05 < \underline{p} < .10$), though as expected the trend was for more able subjects to learn more from context. It is possible that a wider range of ability among subjects would have made the interaction stronger.

Additional analyses were performed to explore the interactions of learning from context with other factors. The interaction of learning from context with Prior Target Word Knowledge was significant for the multiple choice data, $F = 7.58$, $p < 0.01$; subjects learned more about words not previously known. There was no such trend in the interview analysis.

No other interactions with learning from context were found. Notably, the interaction with Level was not significant ($Fs < 1.0$) for either the interview and multiple choice data. Thus, amount of learning from context is independent of the criterion of word knowledge. Other variables that did not influence learning from context were the sex of the subject, the sex of interviewer, the interaction of subject's and interviewer's sex, standardized vocabulary scores, general vocabulary knowledge as measured in the checklist test, interviewer identity, interviewer's teaching experience, version of multiple choice test used, the day and session the subject was tested, subject's background knowledge of the passage topic as measured in the checklist test, and order of the target words as they occurred in the interview.

Discussion

Our results make the important demonstration that learning from context does take place. While the context effect was small in absolute terms, it was statistically robust and very consistent across types of text, methods of measurement, and

levels of scoring. There can be no doubt that the effect was real.

The finding that children do learn word meanings from context is noteworthy because of the materials that were employed: The texts were natural texts, and the contexts were natural contexts. Of the 30 target words, 23 occurred only once. The contexts, especially in the narrative, were not very informative.

The amount of learning from the narrative was the same as that from the exposition. A sample of two texts could hardly be taken as representative of their respective genres; but it is worth emphasizing that the learning of word meanings from context was not confined to the exposition, which, of course, was intended to introduce and explain concepts the author assumed the reader would not know.

Comparison of Findings with Those of Other Research

One way to evaluate the reliability of the present results is to compare them with those of other, similar experiments. This is not a big task in this case, since to the best of our knowledge there is only one experiment in the literature that is really directly comparable to ours, the recent one reported by Jenkins, Stein and Wysocki (in press). The basic design of that experiment was similar to this one: Subjects read texts containing difficult target words and were then tested on their knowledge of these words on several measures. The two studies

did differ, however, in a number of respects that could have influenced the results.

There are some ways in which the experiment by Jenkins, Stein, and Wysocki might have been more conducive to learning from context than ours. One is that, although natural in style, the texts were deliberately written to be informative about target word meanings: "The paragraph context strongly implied the meaning of the target word, and in most cases contained a synonym for the target word (e.g., argument for altercation) in addition to other types of context clues (e.g., temporal, spatial, descriptive)." Contexts meeting these criteria will be richer on the average than the ones in the natural texts used in the present study.

Jenkins, Stein, and Wysocki also had subjects undergo "familiarization training" two days before the start of the body of the experiment. "This training consisted of word reading practice, and was accomplished by teacher demonstration followed by unison reading from the board. No mention was made of any word meanings." This treatment probably caused the subjects to pay more attention to the new words in the texts than they otherwise would have. In contrast, in the current study, subjects were exposed to the target words before reading the passage only in the uninformative checklist task administered three days before the main body of the study, in which the target

words constituted only a small percentage of the 186 items in the test.

Another important difference was the number of repetitions. In Jenkins, Stein, and Wysocki's experiment, subjects were exposed to a target word either two, six, or ten times, each time in an informative context. In the current study, on the other hand, only seven of the 30 target words occurred more than once. The Jenkins et al. study was specifically designed to investigate the effects of repetitive exposure to unfamiliar words in context. In the present study, on the other hand, since natural text was used, the number of occurrences of a target word was not manipulated.

The factors just mentioned are reasons why the Jenkins, Stein, and Wysocki study might show more learning from context than the present one, and also might overestimate incidental learning from context during normal free reading. There are also, however, several differences between the two studies which would tend to cause Jenkins et al. to show less learning from context than was observed in this study.

One is the age of the children used as subjects. In our study, the children were eighth grade students tested toward the end of the school year; thus, their average age was a little over thirteen. The subjects used by Jenkins et al. were fifth grade students; their average age was close to ten years. It is

possible that a three year difference in age put our subjects at an advantage in learning words from context.

Werner and Kaplan (1952) studied the ability of children from ages 8 to 13 to derive the meanings of novel words from context. They found improvement on this task with age, with some aspects of performance changing gradually and others showing abrupt shifts. Big shifts in performance occured between 10 and 11 years, that is, roughly during fifth and early sixth grades. Thus, Jenkins et al.'s fifth grade students might not be expected to learn as much from context as our eighth grade students. On the other hand, children are able to learn new words from oral context, at least at a very early age. Keil (1981), testing children in kindergarten and grades 2 and 4, found that even the youngest subjects were able to make inferences about the meanings of new words encountered in context. From common observation, it is obvious that this ability is present in the preschool years as well. While there might be some developmental change between fifth and eighth grades, it is not likely that the ability to learn meanings from written context would undergo its most significant development only after the fifth grade.

Probably the most important difference between the Jenkins, Stein, and Wysocki study and ours is the way word knowledge was measured. In general, we can say that in the Jenkins et al. study, subjects were given credit for knowing a word only if their answer showed complete, adult-like knowledge of the

meaning. In the case of the multiple choice test, the distractors frequently were similar in meaning to the correct answer, often antonyms or other close semantic relatives. Thus the multiple choice items used by Jenkins et al. were comparable in difficulty to our multiple choice items at the third, or highest, level of difficulty.

The Supply Definitions task used by Jenkins, Stein, and Wysocki corresponds approximately to our interview task. In both cases, the subject was required to provide, rather than to choose, the correct meaning for the target item. A comparison of our scoring with theirs indicates that the scoring for their Supply Definitions task is somewhat stricter than that for our third or highest level of difficulty on the interview ratings.

Another factor that might have made our interview easier than Jenkins, Stein, and Wysocki's Supply Definitions task is that in the latter task, subjects were required to write definitions, while in our interviews, subjects were asked to say what the target words meant. Our interview process was specifically designed to obtain information about subjects' word knowledge that might not appear in written definitions. If subjects didn't respond, or gave incomplete or vague answers, interviewers were instructed to probe further to make sure that as much as possible of the subject's knowledge of the word was elicited.

For both types of tasks, then--choosing a correct meaning in a multiple choice test or providing an oral or written explanation--the criteria for word knowledge imposed by Jenkins, Stein, and Wysocki were at least as high as those required by our strictest measures.

One more factor that could have contributed to a difference in results between the two studies is the amount of delay between the time the passages were read and the time word knowledge was tested. In the Jenkins, Stein, and Wysocki study, there was a two-day delay between the subject's last exposure to a word in context and the administration of the posttests. In the case of subjects receiving only two exposures to the target words, there was an nine-day gap between the two exposures as well. In the present study, interviews about the meanings of the target words began about 15-30 minutes after a subject had read the experimental passage. The story memory task performed during this interval also kept the subjects' attention on the text just read, and on the target words as well.

In summary, the task facing Jenkins, Stein, and Wysocki's subjects was less difficult than that facing ours in that the contexts were richer, the words were repeated more often, and the subjects had their attention drawn to the target words by the "familiarization training." On the other hand, it was more difficult in that their subjects were three years younger, the criteria for demonstrating word knowledge were stricter, and

there was a greater delay between the time of reading and the time of testing.

How do Jenkins, Stein, and Wysocki's results compare with ours, then? In one sense, our results give stronger evidence of learning from context: We found clear evidence of learning from context for target words the majority of which occurred only once in truly natural texts. Jenkins et al., on the other hand, embedded words in less natural texts 2, 6, or 10 times, and did not find statistically significant learning from only two exposures. The most noteworthy fact about the two studies, though, is that both did find significant learning from context. In fact, given the differences in the two studies, the amounts of learning are rather similar, when measured in terms of the probability of learning the meaning of an unknown word from context.

Vocabulary Growth Attributable to Learning from Context

What is the probability of a child's learning an unknown word occurring in a natural written context? The present study allows an answer to this question. "Learning a word" can be defined with respect to any of the criteria for word knowledge that were used. The probability of learning a word to a given criterion equals the increase in number of words known to the given criterion divided by the number of words originally not known to that criterion. Because we did not want to alert subjects to the purpose of the experiment before they read the

passages, neither the interviews nor the multiple choice test were given beforehand to determine prior knowledge of the words. Hence, a direct comparison between pretest and posttest knowledge cannot be made. However, the level of knowledge of target words in the passage subjects did not read was determined. This permits a good estimate of the prior knowledge of the subjects who did read a given passage, since the two groups of subjects did not differ on any measure of prior knowledge or ability. Table 7 gives the probability of learning an unfamiliar word to each level of knowledge assessed in the experiment. For example, at the most stringent criterion of what it means to know a word (Interview Level 3), the probability of learning an unknown word from an exposure in context is about .10 or .11.

Insert Table 7 about here.

It is also possible to derive estimates of the probability of learning a word from context from the results of Jenkins et al. They do not have data for learning from one exposure. However, probabilities of learning a word from 2, 6 and 10 exposures can be calculated from their results. Probabilities for one exposure can then be estimated, assuming the following relationship:

$$P_n = 1 - (1 - P_1)^n$$

In this equation P_n is the probability of learning a word from context on the basis of n exposures; P_1 is the probability of learning a word on the basis of one exposure. The probabilities based on Jenkins et al.'s results are given in Table 8.

Insert Table 8 about here.

Note that the one-exposure probability calculated from the 10-exposure data is less than that calculated from 2- or 6-exposure data. This suggests that the formula above did not satisfactorily compensate for diminishing returns from later exposures. Therefore, the higher one-exposure figure is likely to be more accurate.

The similarity between the probabilities based on our results and those of Jenkins, Stein, and Wysocki is gratifying. According to Jenkins et al.'s data, the probability of learning a word from context to the point of being able to correctly answer a multiple choice question is about .10. From our results, the probability of learning a word to the criterion of Multiple Choice Level 3 (the level closest to Jenkins et al.'s multiple choice criterion) is about .15. The odds of learning a word from a single exposure in context to the point of being able to provide a complete and accurate definition are .05, based on Jenkins et al.'s results, and .11 based on ours. The younger age of the subjects, stricter criteria for word knowledge, and

greater delay between reading and testing in the Jenkins et al. study could easily account for differences of this size.

The picture is somewhat complicated by the fact that the contexts in the Jenkins, Stein, and Wysocki study were richer than those in ours. Further experimentation is necessary to determine how large the effects of the various factors distinguishing these experiments are. In the meantime, we feel fairly confident in assuming that the true probability of learning an unknown word from one exposure in context lies somewhere in the range defined by our study and Jenkins et al.'s.

These probabilities may seem low; but an accurate assessment of its magnitude depends on how many unknown words a child encounters in context during a year. For example, if a child were to encounter 10,000 unknown words, he or she might learn 1000-1500 of them well enough to get the right answer on a multiple choice vocabulary test.

How many unknown words does a child encounter in a year? Unfortunately, information on this point is very scanty, so the best we can offer are tentative estimates. First one needs to have an idea of the total volume of reading. Wilson, Fielding and Anderson (in preparation) have asked fifth grader students to complete daily logs of out-of-school activities over periods ranging from two to six months. From measures of reading speed and minutes spent in reading per day, the yearly volume of exposure to printed language was estimated. A wide range was

found, with some children reporting no reading outside of school, and others reading over 4 million words a year. The median volume of reading is around 700,000 words per year. Since this research was confined to out-of-school reading, it seems safe to estimate that the average fifth grade student encounters more than a million words of written text a year.

How many of these words are unknown? From the present study, we know that the numbers of target words not known in the approximately 1,000-word experimental texts were 8, 11, and 13 at interview levels one, two, and three, respectively. These numbers reflect the number of unknown target words. The target words consisted of the fifteen most difficult words from each text, but the texts also contained other potentially difficult words, some of which were certainly not known by many subjects. The foregoing numbers are therefore underestimates of the total number of unknown words per 1000 words of text. Furthermore, while the texts were appropriate for eighth grader students, the students were above average in ability. This would also decrease the number of unknown words in our results.

Anderson and Freebody (unpublished, but see 1983) have made the most ambitious attempt to date to estimate numbers of unknown words per 1000 words of text. From their research, it appears that a 50th percentile fifth grade student would not know 30 of the words in an average 1,000-word text at even a lenient

criterion of word knowledge, and would not know 59 words at a strict criterion of word knowledge.

In summary, then, according to the best available evidence, (a) the odds that a child in the middle grades will acquire a full adult understanding of an unknown word as a result of one exposure in a natural context may lie between .05 and .11, (b) the number of unknown words that the middle grade child encounters in a representative 1000 word text is between 15 and 55, and (c) the number of words the average middle grade child encounters in print in a year is about 1,000,000. Putting these figures together, the number of new words the typical middle grade child learns in a year from context during reading is between 750 and 5,500; the point-value estimate is 3,125.

The foregoing figures assume a test in which the student must construct answers. However, investigators estimating total year to year vocabulary growth have generally used multiple choice tests. Whereas there is good reason to be distrustful of the validity of multiple choice tests (see Anderson & Freebody, 1983), there is nothing we can do about the preferences of previous investigators. For the purpose of comparison, therefore, we must use Multiple Choice Level 3 as the criterion of word knowledge. Our results show that the probability of learning a word from context to this criterion is between .10 and .15. Thus, if the multiple choice test criterion were accepted as valid, the lower- and upper-bound estimates of annual

vocabulary growth attributable to learning from context would be 1,500 and 8,250; the point-value estimate would be 4,875.

How do these figures compare with children's actual vocabulary growth? There is regretably little consistency among different researchers' estimates of children's absolute volcabulary size (Anderson & Freebody, 1981), and hence wide variation in estimates of yearly vocabulary growth as well. Differences among estimates can be traced to three major sources: The definition of "word" used, the corpus or dictionary used to estimate the total word stock of the language, and the criterion for word knowledge. Nagy and Herman (in preparation), recalibrating earlier estimates to adjust for the first two of these differences, found that adjusted estimates of yearly vocabulary growth converged to a range between 2,000 and 3,600 words a year, with a median figure around 3,300. Comparing this figure with the estimates of yearly learning from context, it appears that incidental learning from written context can account for a large proportion of a child's vocabulary growth during the school years.

There are two types of limitations on the extrapolations we have made from our results. First, there are limitations inherent within the study itself. For example, the short interval between reading and testing may have lead to an overly optimistic assessment of the amount of learning from context. Similarly, the story memory task between reading and testing also

kept the subjects' attention on both the text and the target words, thus possibly improving their performance. The fact that the subjects were eighth grade students and all able readers puts some limits on the generalizability of our results, as does the small number of texts used and the limited number of target words. The similarity between our results and those of Jenkins, Stein and Wysocki does increase confidence in the conclusions, however.

Another limitation on our extrapolations stems from lack of reliable information about factors such as amount of reading done by school children and the number of unknown words they encounter in text. Nonetheless, we believe the figures we used are plausible and fairly conservative; thus we are confident in the general order of magnitude of the estimates. Despite the uncertainties, our analysis suggests that words learned incidentally from context are likely to constitute a substantial proportion of children's yearly vocabulary growth.

Comparison with Direct Vocabulary Instruction

Earlier research gave reason to question the efficacy of learning words from context. The current study shows that the relative value attributed to learning from context and other more direct forms of vocabulary instruction depends largely on the way in which the comparison is made. Our results call to mind the fable of the tortoise and the hare. For any given small set of words, it is easy to show that direct vocabulary instruction is

superior to learning from context. It would be a poor method of instruction indeed that gave a student only a one-in-ten chance of learning an instructed word! But if one asks a different question--what approach to vocabulary can more effectively lead to the acquisition of several thousand words per year--our results indicate that learning from context would be an easy winner. Instruction dealing with words one at a time simply cannot cover that much ground.

Approaches to vocabulary acquisition might be evaluated in terms of time spent per word learned. The intensive vocabulary instruction program implemented by Beck and her colleagues (Beck, McCaslin & McKeown, 1980; McKeown, Beck, Omanson & Perfetti, 1983) is very expensive in this respect. If one divides the increase in number of words known by the total instructional time, an average of .02 words are learned per minute of instruction. In contrast, using the Multiple Choice Level 3 criterion of knowledge (the one most similar to Beck's criterion), about .25 words were learned per minute in the current study. Beck and her colleagues are working on more time-efficient methods of instruction, and in any case, comparisons of rate of learning are fraught with difficulties. Still, it does seem that the impression that direct vocabulary instruction is more efficient than learning from context is an illusion.

Any comparison of approaches ought to take account of the fact that time spent in reading has more benefits than just

growth in vocabulary -- for example, pleasure, gains in general knowledge, and practice in various reading subskills. No doubt the ancillary benefits of direct vocabulary instruction are less rich.

Other Findings

An auxiliary hypothesis investigated in the present research was that good readers would have a higher likelihood of learning word meanings from context than poor readers. The interaction of reading comprehension test performance and contextual learning was not significant, although there was a trend in the expected direction. These results may be attributable to the fact that the range of reading ability was restricted--only average and above average readers participated--and the fact that the standardized test of reading ability was too easy, with many subjects scoring near the ceiling.

The fact that learning from context takes place at all levels of word knowledge means that context is not limited to providing only a vague, initial indication of a word's meaning. Although contexts that precisely identify a word's meaning may be relatively rare, our results show that many contexts provide enough information to help the reader reach a full adult understanding of the meaning of a word.

On the other hand, our results are still consistent with a model in which the learning of individual word meanings proceeds in terms of small increments. The subjects knew about half of

the target words from the passage they had not read at at the level of Interview Level 1, and Multiple Choice Levels 1 and 2. So it is very likely that words learned from context to higher criteria of knowledge were already partially known. Our results also agree with the widely-noted fact that children's vocabularies contain large numbers of partially-known words.

Conclusion

The major result of our study has been to demonstrate unmistakable learning from context from one or a very few exposures to unfamiliar words in natural text. This finding will not surprise those who have believed all along that learning from written context is a major source of vocabulary acquisition. It is surprising considering that previous experimental studies often have failed to find significant learning from context, even studies that used contrived contexts richer than the ones typical in nature. The showing that learning from context makes vis-a-vis other methods of vocabulary learning depends on how the comparison is made; the strength of learning from context lies in its long-term, cumulative effects.

The present study was concerned exclusively with written contexts. Oral contexts also play a major role in vocabulary growth. Indeed, the importance of exposure to vocabulary in rich oral contexts cannot be overestimated, particularly for young children. But large areas of a student's oral language environment--the speech of parents and peers--are mostly beyond a

teacher's control. Our results, on the other hand, suggest that a moderate amount of reading, which a teacher can influence, will lead to substantial vocabulary gains. Furthermore, in terms of words learned per minute, learning from context is likely to compare favorably with direct vocabulary instruction, which is the other alternative a teacher has.

We would not care to maintain that no direct instruction in vocabulary should ever be undertaken. But, as we have argued elsewhere (Nagy & Anderson, 1984; Nagy & Herman, in preparation), the number of words to be learned is too enormous to rely on word-by-word instruction. It follows that students must somehow become independent word learners. So far, attempts to design direct vocabulary instruction that generalizes, leading students to independently learn non-instructed words, have failed (cf. McKeown, Beck, Omanson, & Perfetti, 1983). On the other hand, our results strongly suggest that a most effective way to produce large-scale vocabulary growth is through an activity that is all too often interrupted in the process of reading instruction: Reading.

References

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), Comprehension and teaching: Research reviews (pp. 77-117). Newark, DE: International Reading Association.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.), Advances in reading/language research (pp. 231-256). Greenwich, CT: JAI Press.

Arthur, R. (1981). The midnight visitor. Beacons. New York: Houghton-Mifflin. (Originally published from Mystery and more mystery. New York: Random House, Inc., 1967)

Beck, I., McCaslin, E., & McKeown, M. (1980). The rationale and design of a program to teach vocabulary to fourth-grade students (LRCD Publication 1980/25). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Beck, I., McKeown, M., & McCaslin, F. (1983). All contexts are not created equal. Elementary School Journal, 83, 177-181.

Beck, I., McKeown, M., McCaslin, F., & Burkes, A. (1979). Instructional dimensions that may affect reading comprehension: Examples from two commercial reading programs (LRCD Publication 1979/20). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Bishop, M., Sutherland, B., & Lewis, P. (1981). Water systems. Earth science. Columbus, OH: Charles Merrill Publishing Company.

Boettcher, J. (1980). Fluent readers' strategies for assigning meaning to unfamiliar words in context. Unpublished doctoral dissertation, University of Minnesota. (University Microfilms International: Catalog No. 80-19516).

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), Linguistic theory and psychological reality (pp. 264-293). Cambridge: MIT Press.

Carey, S. (1982). Semantic development: The state of the art. In E. Wanner & L. Gleitman (Eds.), Language acquisition: The state of the art (pp. 345-389). Cambridge: Cambridge University Press.

Carroll, J., Davies, P., & Richman, B. (1971). The American Heritage word frequency book. Boston: Houghton-Mifflin.

Carroll, B., & Drum, P. (1983). Definitional gains for explicit and implicit context clues. In J. Niles & L. Harris (Eds.), Searches for meaning in reading/language processing and instruction (pp. 158-162). Rochester, NY: The National Reading Conference.

Clark, E. (1973). What's in a word: On the child's acquisition of semantics in his first language. In T. Moore (Ed.), Cognitive development and the acquisition of language (pp. 65-110). New York: Academic Press.

Dale, E., O'Rourke, J., & Bamman, H. (1971). Techniques in teaching vocabulary. Palo Alto, CA: Field Educational Publications.

Deighton, L. (1959). Vocabulary development in the classroom. New York: Bureau of Publications, Teacher's College, Columbia University.

Dupuy, H. P. (1974). The rationale, development and standardization of a basic word vocabulary test. Washington, DC: U.S. Government Printing Office. (DHEW Publications No. HRA 74-1334)

Durkin, D. (1979). What classroom observations reveal about reading instruction. Reading Research Quarterly, 14, 481-533.

Eichholz, G., & Barbe, R. (1961). An experiment in vocabulary development. Educational Research Bulletin, 40, 1-7, 28.

Freebody, P., & Anderson, R. (1983). Effects on text comprehension of differing proportions and locations of difficult vocabulary. Journal of Reading Behavior, 15, 19-39.

Gentner, D. (1975). Evidence for the psychological reality of semantic components: The verbs of possession. In D. A. Norman & D. E. Rumelhart (Eds.), Explorations in cognition (pp. 211-246). San Francisco: W. H. Freeman & Co.

Gibbons, H. (1940). The ability of college freshman to construct the meaning of a strange word from the context in which it appears. Journal of Experimental Education, 9, 29-33.

Gipe, J. (1979). Investigating techniques for teaching word meanings. Reading Research Quarterly, 14, 624-644.

Gipe, J. (1981). Investigation of techniques for teaching new words. Paper presented at the Annual meeting of the American Educational Research Association, Los Angeles.

Gray, W., & Holmes, F. (1938). The development of meaning vocabularies in reading (No. 6). Chicago, IL: University of Chicago Publications of Lab Schools.

Jenkins, J., & Dixon, R. (1983). Vocabulary learning. Contemporary Educational Psychology, 8, 237-260.

Jenkins, J., Pany, D., & Schreck, J. (1978). Vocabulary and reading comprehension: Instructional effects (Tech. Rep. No. 100). Urbana: University of Illinois, Center for the Study of Reading.

Jenkins, J., Stein, M., & Wysocki, K. (in press). Learning vocabulary through reading. American Educational Research Journal.

Johnson, D., Toms-Bronowski, S., & Pittelman, S. (in press). The investigation of the effectiveness of semantic mapping and semantic feature analysis on vocabulary acquisition and retention. Reading Research Quarterly.

Keil, F. (1981). Semantic inferences and the acquisition of word meaning. Unpublished manuscript, available from F. C. Keil, Dept. of Psychology, Cornell University, Ithaca, NY 14853.

Margosein, C., Pascarella, E., & Pflaum, S. (1982). The effects of instruction using semantic mapping on vocabulary and comprehension. Paper presented at the annual meeting of the American Educational Research Association, New York.

McKeown, M., Beck, I., Omanson, R., & Perfetti, C. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. Journal of Reading Behavior, 15, 3-18.

Nagy, W., & Anderson, R. (1984). The number of words in printed school English. Reading Research Quarterly, 19, 304-330.

Nagy, W., & Herman, P. (in review). Limitations of vocabulary instruction. Urbana: University of Illinois.

Rankin, E., & Overholzer, B. (1969). Reaction of intermediate grade children to contextual clues. Journal of Reading Behavior, 2, 50-73.

Sachs, H. (1943). The reading method of acquiring vocabulary. Journal of Educational Research, 36, 457-464.

Smith, M. (1941). Measurement of the size of general English vocabulary through the elementary grades and high school. General Psychological Monographs, 24, 311-345.

Sternberg, R., & Powell, J. (1983). Comprehending verbal comprehension. American Psychologist, 38, 878-893.

Templin, M. (1957). Certain language skills in children: Their development and interrelationships. (The Institute of Child Welfare Monographs, No. 26). Minneapolis: University of Minnesota.

Werner, H., & Kaplan, E. (1952). The acquisition of word meanings: A developmental study. Monographs of the Society for Research in Child Development. (15, Serial No. 51).

Wilson, P., Fielding, L., & Anderson, R. C. (in preparation). Prolegomena to the study of free reading. Urbana: University of Illinois.

Table 1

Target Words

Narrative (Spy) Passage	Expository (River System) Passage
authentic	bed
confounded	divide
countenance	drainage basin
disillusioned	headward extension
envision	impermeable
espionage	levee
explanatorily	meander
gendarme	oxbow lake
moodily	porous
passably	rill
passkey	runoff
prosaic	saturated
twang	suspended load
wheezily	turbulent

Table 2

Example of Three Levels of Multiple Choice Items

LEVEL ONE

gendarme means:
a) to trick or trap someone
b) policeman
c) spoken as if one was out of breath or having trouble breathing
d) the secret collection of information about another country
e) the illegal transportation of goods across a border
f) don't know

LEVEL TWO

gendarme means:
a) the illegal transportation of goods across a border
b) weapon
c) policeman
d) face
e) bravery during wartime
f) don't know

LEVEL THREE

gendarme means:
a) policeman
b) bellboy
c) bodyguard
d) spy
e) waiter
f) don't know

Table 3

Examples of Levels of Word Knowledge in Interview Scoring:

Attempts to Define the Word "Disillusioned"

Student Answer	Score
"not illustrated correctly"	0: no correct knowledge
"I think it's something imagined . . . a picture of something in your mind"	1: answer shows mental activity and is vague
"If you're like led astray. If you're made to believe something that's not really true."	2: answer does not convey that the person must realize the deception and consequently feel let down and disappointed
"If you have ideas about something and you find out it's the opposite, you're disillusioned. Your beliefs are shattered."	3: answer conveys a complete understanding

Table 4

Percentage of Words Known at Each Criterion Level

Level of Word Knowledge	Measure of Word Knowledge					
	Interview			Multiple Choice*		
	Text Read	Wordsource		Text Read	Wordsource	
		Narrative	Exposition		Narrative	Exposition
Level One	Narrative	58	47	Narrative	70	52
	Exposition	48	57	Exposition	64	59
Level Two	Narrative	41	21	Narrative	64	51
	Exposition	30	32	Exposition	57	64
Level Three	Narrative	21	10	Narrative	59	37
	Exposition	12	19	Exposition	54	47

*Multiple choice scores are corrected for guessing

Table 5

Analyses of Interview Data

Variable	B	% Variance	F
Subject's Grand Mean	0.9	7.7	501.7
Prior Target Word Knowledge[a]	8.5	2.8	185.4
Level[b]	-18.6	10.5	683.6
Text Read[c]	- 0.1	0.0	0.0
Word Source[c]	3.1	0.6	36.3
Learning from Context[d]	- 3.4	1.2	75.8
Comprehension	0.2	0.0	0.3
Comprehension x Learning from Context	0.8	0.1	4.7
Constant/Residual	36.7	77.2	

Note. Critical value (1,5049) = 6.85, $p < .01$

[a] Coded 1,0

[b] Coded 1,2,3

[c] Coded +1 narrative; -1 expository

[d] Coded +1 words from passage read; -1 words from passage not read

Table 6

Analysis of Multiple-Choice Data

Variable	B	% Variance	F
Subject's Grand Mean	0.9	5.7	332.2
Prior Target Word Knowledge[a]	10.9	3.3	192.3
Level[b]	- 6.0	0.9	49.6
Position[d]	3.0	0.9	51.5
Text Read[c]	- 0.5	0.0	0.2
Word Source[c]	6.3	1.5	89.4
Learning from Context[e]	- 4.3	0.6	34.3
Comprehension	0.4	0.0	0.9
Comprehension x Learning from Context	0.8	0.0	3.3
Constant/Residual	0.5	87.0	

Note. Critical value (1,5046) = 6.85, $p < .01$

[a] Coded 1,0

[b] Coded 1,2,3

[c] Coded +1 narrative; -1 expository

[d] Coded 1-6 (blocks)

[e] Coded +1 words from passage read; -1 words from passage not read

Table 7

Probability of an Unknown Word Being Learned to a Given Criterion Level

	Wordsource		
	Narrative	Expository	Mean
INTERVIEW			
Level 1	.194	.192	.193
Level 2	.160	.139	.150
Level 3	.110	.102	.106
MULTIPLE CHOICE			
Level 1	.194	.205	.200
Level 2	.187	.249	.218
Level 3	.154	.148	.151

Table 8

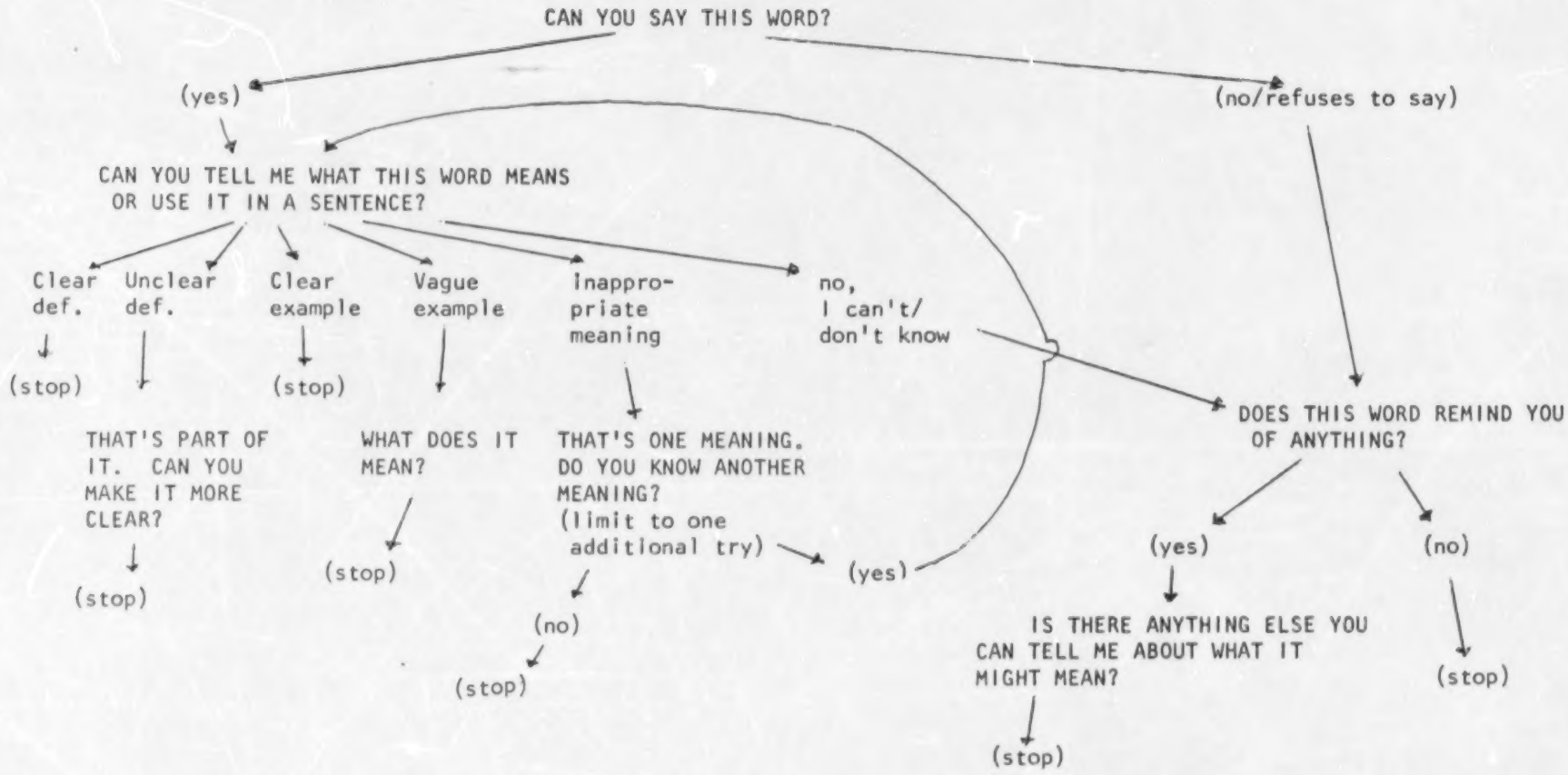
Probabilities of Learning Word From Context Based on Results from Jenkins, Stein & Wysocki (in press)

	Number of Exposures to Word In Context	Probability of Learning Word From Total Exposures	Probability of Learning Word From One Exposure
SUPPLY DEFINITION TASK	2	.101	.052
	6	.285	.054
	10	.196	.022
MULTIPLE CHOICE	2	.198	.104
	6	.362	.072
	10	.453	.059

Figure Caption

Figure 1. Flowchart for Interviewers' Use of Prompts.

INTERVIEW FORMAT*



*This gives the general strategy of interviewing and the depth of probing required. Wording can be changed to suit experimenter, and what seems to work for subject. Also, repetition of questions may be unnecessary as subject becomes familiar with procedure--as long as interviewer remembers to probe when necessary.